NBSIR 85-3154-1

Applied Model Validation

Alan D. Davies

U.S. DEPARTMENT OF COMMERCE
National Bureau of Standards
National Engineering Laboratory
Center for Applied Mathematics
Gaithersburg, MD 20899

Revised July 1985

Sponsored in part by
U.S. DEPARTMENT OF COMMERCE
National Bureau of Standards
er for Fire Research
ersburg, MD 20899

NBSIR 85-3154-1

APPLIED MODEL VALIDATION

Alan D. Davies

U.S. DEPARTMENT OF COMMERCE
National Bureau of Standards
National Engineering Laboratory
Center for Applied Mathematics
Gaithersburg, MD 20899

Revised July 1985

Sponsored in part by
U.S. DEPARTMENT OF COMMERCE
National Bureau of Standards
Center for Fire Research
Gaithersburg, MD 20899

U.S. DEPARTMENT OF COMMERCE, Malcolm Baldrige, *Secretary*
NATIONAL BUREAU OF STANDARDS, Ernest Ambler, *Director*

Table of Contents

List of Figures

ABSTRACT

This is a progress report of an applied model validation case study. The subject model is "Transport of Fire, Smoke and Toxic Gases (FAST)" by W. W. Jones of the National Bureau of Standards Center for Fire Research. Products from a fire in a "burn room" exit through a connected corridor to outdoors. Cooler counterflow air in a lower layer feeds the fire. The model predicts corridor layer temperatures and thicknessess vs. time, given enclosure, fire and ambient specifications. Data have been collected from 38 tests using several fire sizes. Corresponding model results, and model and test documentation are yet to come.

Considerable modeling and calculation is needed to convert instrument readings to test results comparable with model outputs so that residual differences may be determined. Test results as well as model results must be validated, and test result uncertainties estimated so that they are not unfairly attributed to the model.

APPLIED MODEL VALIDATION REPORT--A PROGRESS REPORT

1. INTRODUCTION

The NBS Center for Fire Research (CFR) conducts scientific research bearing on the fire safety of buildings, vehicles, tunnels and other inhabited structures. Similar and related work has been done over many years at Harvard University, BRI from Japan, California Institute of Technology and the University of Dayton as well as elsewhere. Data from controlled fire experiments are collected, analyzed and reduced to the analytical formulas that appear to underly the observed phenomena. These results and more general physical principles are then combined into models to predict the development of environments that may be hostile to humans. The tests provide points of departure for much less expensive exploration using the model. A major purpose of such work is to help fire safety officials in the appraisal of existing and proposed buildings and vehicles. The safety of people, rather than property, is the point.

Building escape and evacuation models exist that use assumptions about occupancy, escape routes, and link lengths, capacities and movement speeds to estimate escape times. References [3]*, [5] and [16] are representative of such work. Comparing these times with the time profiles for development of the hostile environment provide a basis for assessing fire safety -"Is there enough time to escape?".

Following the NBS policy of cooperation among its organizational units, the Center for Applied Mathematics (CAM) joined with CFR to conduct a formal validation of a selected fire model. The main purpose of this work was to gain practical experience in the validation process, as expounded in a voluminous literature on the subject. References [6], [7], [8], [9], [10], and [15] are indicative. Of course, the model selected for study benefits directly.

This progress report is on a case study of applying a validation process to a selected model. Realism is provided by treating a particular fire effects model and its experimental reference. However, the real subject is the validation process, not the model or the tests. This leads to a two-tier report of a project within a project.

The model is "Transport of Fire, Smoke and Toxic Gases (FAST)" [12] by W.W. Jones of the Fire Safety Technology Division of CFR. It has been under development and improvement since the fall of 1982. References [11], [13], [14] and [17] are a small sample of related work. An early version of the model and its associated computer program was applied to fires in a multi-room structure at NBS. The model has a core of fundamental principles and components that can be reassembled to cover a wide range of cases. Eventually, coverage of surface ships and confined environments such as aircraft and submarines is planned.

* Numbers in brackets, [], refer to citations given in Section 8 of this report.

Detailed plans for a simpler and better instrumented full-scale test facility began in the fall of 1982 under the direction of S. Davis of CFR's Fire Measurement and Research Division. Other principals from the Fire Measurement and Research Division include B.T. Lee, J.S. Steel, R.D. Peacock and J.N. Breese. The first experiment in this new facility was conducted in September of 1983 and 37 more were conducted by the end of March 1984.

CAM involvement began in June 1983 and continues. The CAM role is mainly to advise and help in experimental design (i.e. randomization of the controlled experimental parameters), validation methods and analysis of the data. Detailed suggestions led to further instrumentation, collection of additional data, revision of some of the data reduction programs and of the model's program. CAM personnel were A.D. Davies of the Operations Research Division and J.J. Filliben of the Statistical Engineering Division. Filliben is the author of DATAPLOT [4], a software package used in part of this work.

Although much was learned, major tasks remain before useful validation statements are possible. Nevertheless, a progress report is considered appropriate.

Section 2 contains some validation background and rationales that were used. A description of the model is in Section 3. Section 4 covers the experimental reference, including the test facility, test procedures, data recording, processing, and interpretation. Validation activities are reported in Section 5, some observations in Section 6 and recommendations are in Section 7.

2. BACKGROUND AND RATIONALES

Model validation is essentially a reviewing or auditing function with a bias towards the user's viewpoint. To be convincing, the auditor should be independent of the auditee, although the auditee provides most of the information used by the auditor. For working purposes, the user is assumed to be a building safety official with questions such as:

Will this proposed new building be safe enough, and if not, how might the design be improved?

Is this old building safe enough, given current or proposed new uses?

The main user questions about the model deal with what the model covers (and doesn't cover), how uncertain are the results, and what it takes to use the model. These questions expand into many details that will be grouped under the general headings of user validation and technical validation.

2.1 User Validation

"User" validation examines the instructions and other items that make the model accessible and competitive. Descriptions are needed of the model, its supporting program, input preparation, program operation and maintenance, the computational environment, and values offered for some of the inputs. User validation remains to be done.

Accessibility includes portability issues such as computer hardware, software and programming language requirements, and the practicality and ease of assembling the necessary input data. Relative attractiveness among models involves tradeoffs, some of which may be subjective. Quality of documentation, costs for collecting the data and running the model, the range of cases left unresolved and display features are among items affecting attractiveness.

2.2 Technical Validation

The technical side of validation mainly involves accuracy and uncertainties. The most visible products in this category are statements about systematic and random "residual differences" over time between test and model estimates of the same parameters.

A residual difference is much like a surveyor's "error of closure" around some closed path. The common starting point for the model and the test consists of the experimental conditions, including the facility, initial conditions and relevant test variables. Some result variable such as the height of the interface between hot and cool layers in the test facility is selected for study. Beginning at the common starting point, the result variable is independently estimated around one part of the path using the model and from the other direction using the test. The difference between these two results is the residual difference, or more briefly, the "residual".

The residual will be composed of real differences due to the model and of those due to errors and approximations in the tests and test data interpretation. The result variable is seldom measured directly by a sensor in the test facility. Usually some modeling is required. The validation task is to isolate the component due to the model, which means that both path segments must be studied.

A useful model does not have to resolve every doubt, but the output uncertainties should not leave too many cases for deeper study. Safety officials are concerned with the aggregate of these uncertainties, whatever the source, since these affect the quality of their decisions. This unifies the user's analytical problem and provides a basis for the balanced treatment of the contributing uncertainties. Thus, technical validation focuses repeatedly on the biases and uncertainties that are passed on to the next analytical step.

Confidence in the interpreted experimental (test) results is essential, since errors here propagate into perceived model errors. Hence, the tests must also be closely examined and technically "validated". The test results can then be reused to examine similar models and can become a valuable resource in their own right. As long as these results are not used as reference data for model development, their value for validation purposes is retained. If this independence is lost, a new set of tests would be needed for proper model validation. The tests need to be documented to support residual difference calculations and peer reviews. Most of the work to date has been in technical validation of the test data.

The parameters of primary user concern in this test case are the values over time of gas temperatures and layer thicknesses along the escape route, in or near ranges bearing on human survival. This provides an objective basis and even some rough physical bounds for useful tradeoffs and analyses. Temperatures under which humans can survive are well below those at which combustible wall and ceiling materials burst into flame (called "flashover"). Acceptable escape times are usually a small number of minutes after the fire starts and include the transient period of fire growth.

The error tradeoff aspects are particularly interesting, although they may more properly belong to model and test design rather than validation. Inevitably, errors in some variables in the problem are more influential on the results than others. This applies to variables that arise anywhere in the calculations, not just sensor outputs or model inputs. It also applies to modeling approximations, and to the design, execution and interpretation of the tests. For this reason, technical validation is heavily concerned with error analysis and methods that permit problem subdivision.

Several methods for examining model and test results are available. These range from expert reviews for reasonableness through graphical presentations to formal statistical calculations. The first two methods have been applied successfully. It is too early for formal methods to be very effective.

Graphical methods have proven to be strong. Data streams that should be alike can be developed and compared, deriving streams of residuals. These can then be plotted and examined for random and bias errors, and for systematic behavior called "structure". In other cases, simple time plots were generated and examined for reasonableness. Time is not the only useful independent variable. Any pair of variables can be cross-plotted to see if their mutual behavior makes sense to a knowledgeable reviewer.

Ideally, the residuals will show a uniform random distribution around the zero difference axis. Departures from this ideal then become subjects for explanation, and correction if they are considered serious. Such tests are not limited to comparisons between the test results and the model, but can be devised within parts of the problem if foresight has made the necessary data available. For example, conservation laws for energy and mass flow rates may be used to test data groups for each room and the overall test facility. Combinations of such tests have led to revisions of test data reduction algorithms, thermal constants for the enclosure materials, doorway coefficients of discharge and other features of the data collection and reduction system. Reasonableness tests (e.g. conservation of mass) were also applied to outputs of some early model versions. These failed, indicating problems that have since been addressed. Improvements also continue to a set of CFR general-purpose test data interpretation programs called SPEEDY.

Objective estimates of testing uncertainties are needed to support the fair allocation of uncertainties to the model. Results from tests that are nominally alike will differ for reasons that are beyond the reasonable control of the experimenter. Weather changes, unnoticed changes in the test facility and roundoffs in data recording and computations are among such causes. This

means that the tests need to be repeated (replicated) so that the differences have a chance to occur and be observed. The numbers of replications are compromises between testing costs and the quality of the resulting statistics. Joint CAM/CFR test replication recommendations have been carried out. Consistency checks among nominally identical tests are beginning.

2.3 Follow-on plan

Details of the model, the test facility, testing and the treatment of data will be discussed in later sections of this report and to which the reader is referred for understanding of unfamiliar terms. A general plan for further joint CFR/CAM actions is given below.

a. Review the test data for internal consistency within each fire. At least satisfy heat balances by showing residual errors vs time, and by studying time plots of variables to see if some errors still remain in the SPEEDY modeling.

b. Work out a method for estimating moving standard deviations (msd's) or some equivalent error index. This diagnostic tool is desired so that error problems can be localized to portions of the tests.

c. Develop msd's for the sensor outputs and derived quantities, and indices of uncertainty for the other inputs.

d. Gather and systematize the presentation and input of exogenous parameters.

e. Conduct sensitivity analyses of SPEEDY models with respect to sensor readings and user-supplied inputs.

f. Combine c, d and e to show where accuracy is most important at the sensor level and where relaxation is allowable, or where instrumentation or interpretation models might be rearranged to give a better overall result.

g. Run FAST for each test fire under several conditions:

1. Nominal parameters of ambient conditions and fire sizes and no initial pilot light.
2. Nominal parameters, with initial pilot light (about 3kW).
3. Recorded parameters without pilot light.
4. Recorded parameters with pilot light.

h. Develop and try ways of presenting accuracy or uncertainty statements to potential users. Also identify the acceptable levels of uncertainty in the input parameters, especially where nominal values can suffice. (This may also lead to model simplification.)

i. Prepare testing, model and program documentation and review it for understandability and completeness.

j. Test model portability. Follow the instructions and see if one can get results and how hard it is to do so.

3. MODEL

The model selected for study is "Transport of Fire, Smoke and Toxic Gases (FAST)" by W.W. Jones of CFR. It is programmed in ANSI FORTRAN 77, a language supported on CFR's Perkin-Elmer 3242. This computer has 8 megabytes of random access memory (RAM) and 600 megabytes of disk memory. The model, including the generation of a color display of temperatures and layer thicknesses in each room, occupies about 0.3 megabytes of RAM.

FAST is a "zone" (engineering) model, as distinguished from a "field" (scientific) model. Field models rely on a relatively small number of basic physical laws, examine interactions in more detail than zone models and tend to be used more for studying processes than for answering operational questions. Examples of field models are reported in [1] and [2].

Zone models use lumped constants and other types of approximation to produce usable answers at the potential expense of scientific completeness. For example, this model assumes instant transport and uniform mixing of gases within a gas layer in a room. This means that gas temperatures at a given height are independent of horizontal location in a room, in spite of knowing that temperatures rise somewhat towards the fire. This and other approximations are made deliberately to gain speed and ease of model use.

The purpose of the model is to describe the time-dependent development of potentially threatening toxic gas and temperature conditions in a multi-compartment enclosure. Inputs describe the enclosure, the fire and environmental conditions. The principal outputs are time-varying average upper and lower layer gas temperatures in the sequence of connected rooms in the enclosure and the heights of the interfaces between the hot and cool layers. There are several other parameters generated by the model that can be made available for additional comparisons with test results. These can be useful in validating components of the main model so that potential problems can be isolated and confidence in these submodels can be gained.

Figure 1 is a side view schematic diagram of some elements in the analysis. The room containing the fire is called the "burn room". The burn room is connected through a door to another room called the "corridor". The corridor is vented to outside conditions through a normal door. (A plan view of these rooms will be presented in Section 4.1 on the Test Facilty.)

Two gas layers are shown, a hot upper layer (shaded) and cooler lower layer (clear), separated by a layer boundary. The less dense hot gases are buoyed up by the cooler incoming air. Some mixing (entrainment) occurs across this boundary induced in part by the opposing flows of the two layers. The fire uses incoming air from the lower layer, adding heat and combustion products. The fire plume delivers hot gases and smoke to the upper layer. The hot gases spill out into the corridor once they have filled the burn room down to the top of the burn room doorway. Corresponding hot and cool layers form eventually in each of the connected rooms.

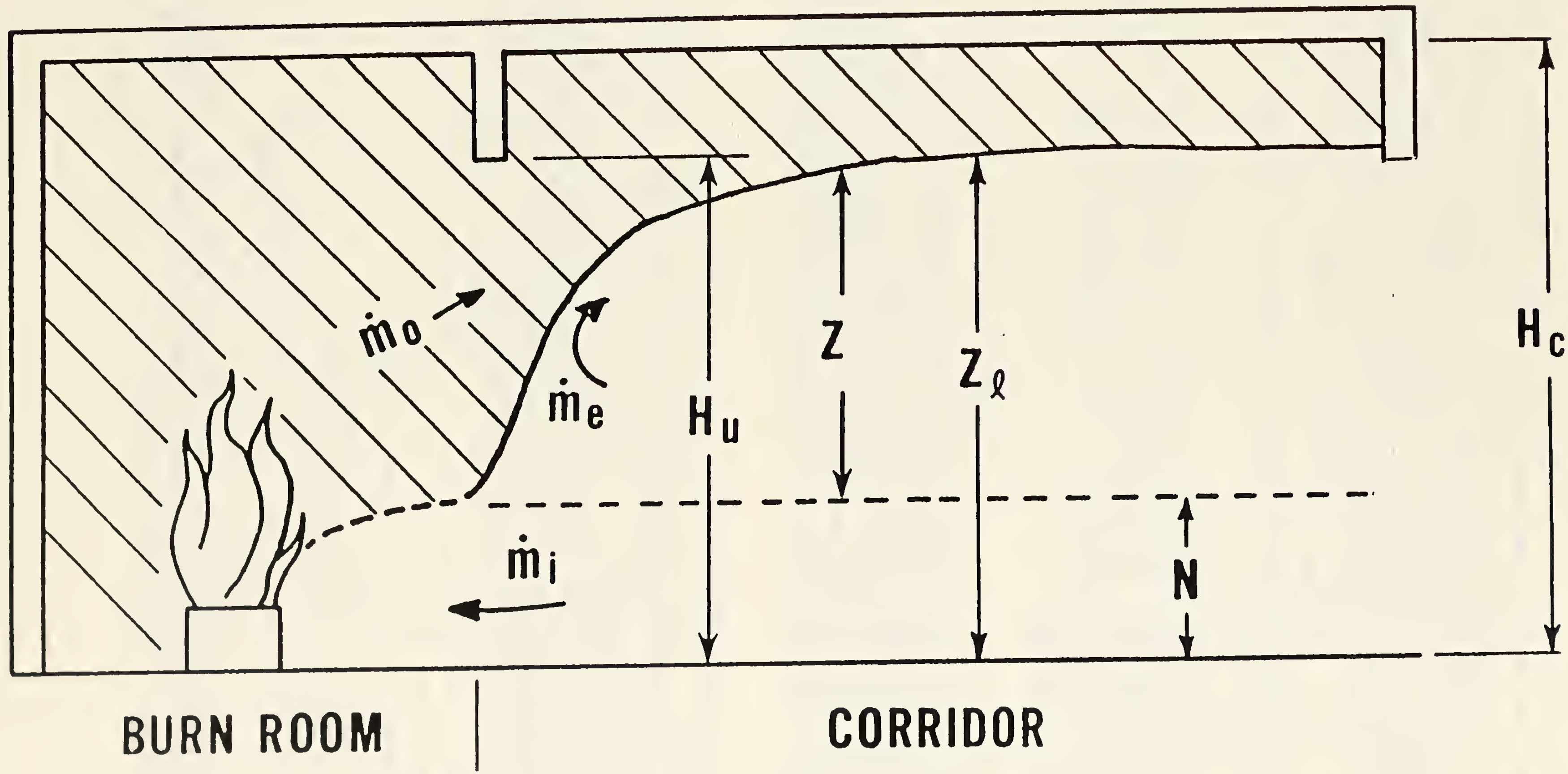

FIGURE 1. TEST FACILITY SCHEMATIC
(SIDE VIEW)

The enclosure shape is described by the inside (rectangular) dimensions of the rooms, and the connections among the rooms and to the outdoors. Heat exchanges with the enclosure are important in the analyses, requiring careful specification of the thermal properties of the enclosing materials. Floors, ceilings and walls can be specified separately. Specific heat, density, emissivity and thermal conductivity of the materials are the relevant items. The thicknesses and sequence of material layers can complicate the enclosure heat absorbtion models, but some simplifying assumptions are used, based on practical relationships between escape times and heat pulse penetration times. Both the model and the data interpretation programs assume semi-infinite enclosure material thicknesses.

The fire is described by location in the enclosure and by its heating rate as a function of time. Heating rates have the dimensions of power, making the use of kiloWatts, kW, convenient. The model has the capability of accepting arbitrary heating rate profiles. Controlled step function inputs were used in the validation tests, but more complicated profiles such as from a mattress fire can be used.

Boundary conditions include the ambient pressure and temperature of the outside air and initial temperatures inside the enclosure. Specifications of the enclosure, the fire and boundary conditions are held in common by the model and the tests. Normally they would be supplied by a model user who may get some of them from architectural specifications, model documentation or even default values supplied as part of the model package. Where validation is being conducted, they come from the test records and comprise a statement of test conditions.

The model inputs include:

o A geometric and thermal description of the enclosure and its apertures.

o The heat release profile of a fire in one of the rooms. The present set of tests began with a baseline period of data collection, then a five-minute burn of a pilot light at about three kW, followed by ignition of the combined main burner and smoke generator. This step function fire was maintained for ten to fifteen minutes and then shut off. The heating profile is the primary independent variable for both the model and the tests. Minor heat contributions come from lights used in the optical density instrumentation and floodlights for camera and visual observations.

o Location of the fire in the burn room.

o Temperature and pressure of the external environment.

As mentioned above, the main outputs of the model are time profiles of average upper and lower gas temperatures within compartments along the escape route and the heights of the layer interfaces.

4. THE EXPERIMENTAL REFERENCE

The purpose of this section is to describe the basic and derived test results to be compared with model results. The goals are to produce test results that agree in concept with the model products within the practical constraints of experimentation, to produce unbiased test values and to quantify the associated uncertainties. Errors and uncertainties are major topics since they affect the value of the model to a user.

Sensors are necessarily at discrete locations which often are not exactly where the parameter estimates are desired. Therefore, analytical models must be used to produce interpolated or extrapolated values with respect to location, and possibly time as well. If some parameter such as the heat flow rate through part of a doorway is needed, the interpreting model may use data from several sensors and constants from other sources. Such models can become quite involved, and their approximations also contribute to the eventual residual differences between model and test results. Both the measurements and the interpreting models are potential sources of bias and random errors and should be subjected to careful scrutiny.

Data describing test conditions can contribute to errors. These data include ambient air pressure, temperature and humidity, heating value of the fuel, fuel input rates, dimensions of the test rooms, thermal properties of materials, locations of sensors and other items. The gas company will provide the higher heating value of the fuel. However, the lower heating value should be used since the water vapor from combustion is not condensed before it leaves the test facility.

Other errors can arise from assumptions about the smoke source and operation of other heat sources such as workers and lights. Standard temperature and pressure assumptions do not really apply since the facility is at an altitude of about 450 feet, not Sea Level. Sometimes the rooms had not cooled down to outside conditions when the test started. Pretest air flows and layering were observed inside the test facility. Thus, the simple initial conditions that are used in casual descriptions are not really present for the tests, and can be sources of probably minor but unnecessary error.

4.1 Test Facility

Figure 2 is a plan view of the test facility. The nominal vertical cross-section of the burn room and the corridor is eight feet high by eight feet across. The nominal lengths are eight feet for the burn room and 40 feet for the corridor. Actual dimensions are reduced by two inches of a high-temperature insulating material called Kaowool on the walls and ceiling of the burn room and a corresponding half inch of Marinite in the corridor. These materials protect the rooms against damage from repeated fires. The stub corridor between the burn room and the corridor is also lined with Kaowool. The burn room floor is fire brick and the corridor floor is Marinite over a concrete slab.

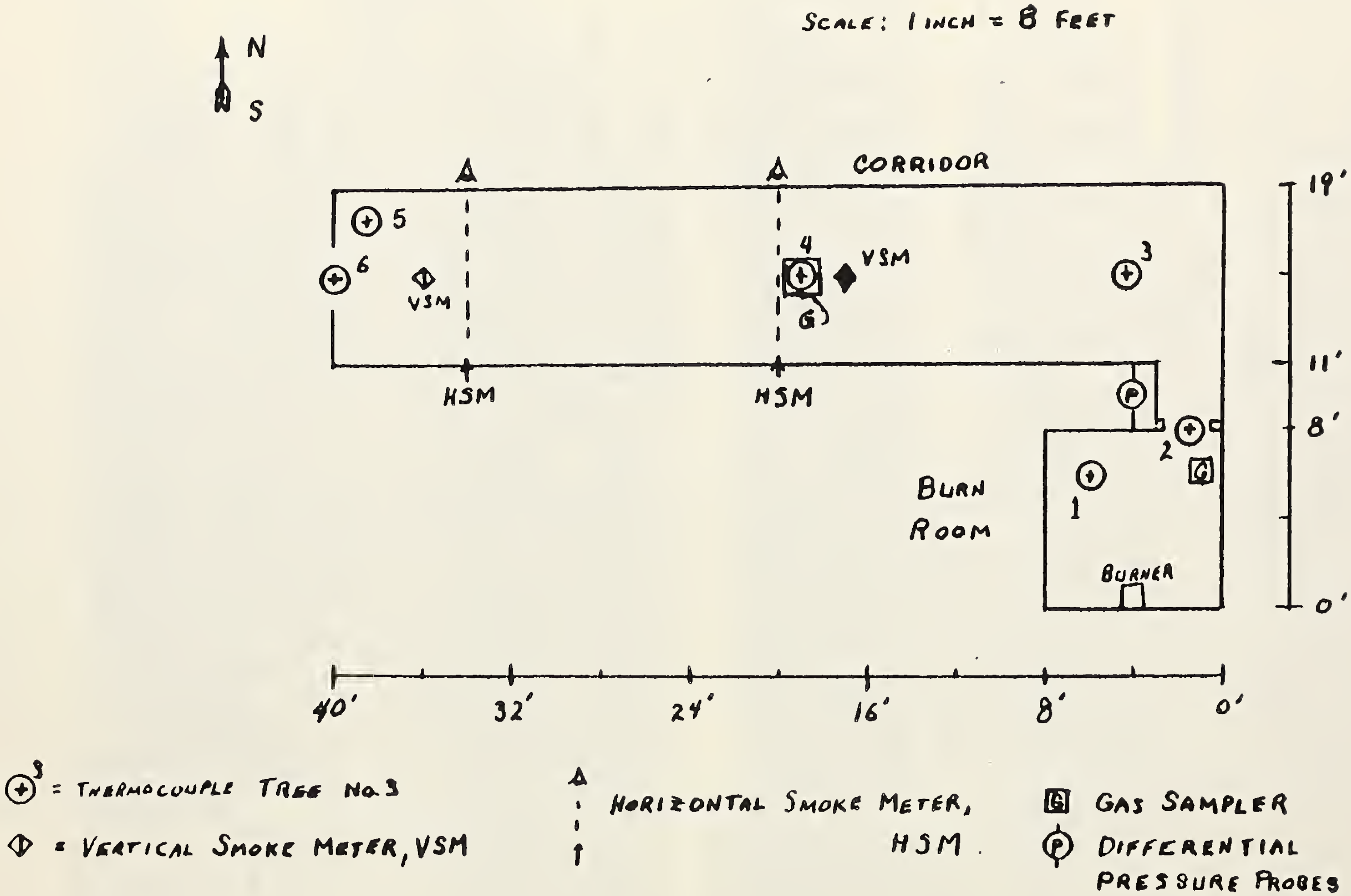

FIGURE 2. TEST FACILITY PLAN VIEW

Locations of several test output sensors are also indicated in Figure 2, including several vertical arrays or "trees" of instruments. Most of the sensors are thermocouples, measuring temperatures. Others measure pressure differences through the walls at doorways (in support of gas flow calculations), total incident radiative and convective heat flux (for heat losses to the enclosure), exit gas oxygen, carbon monoxide and carbon dioxide concentration (for completeness of burn estimates) and attenuation of light intensity (for smoke density).

These output sensors produce voltages that are polled by an automatic data recorder on a ten second cycle at a rate of 24 per second. A digital recording channel with a preselected level of resolution is assigned to each output sensor. Sensor calibration scales are among the data provided to the test data processing program called SPEEDY. There are roughly 100 thermocouples and 50 other test output sensors. These readings are combined with other physical data in the SPEEDY models to compute another 100 channels of derived results.

Sensors that are not automatically sampled and recorded include a barometer, a thermometer and a hygrometer for ambient pressure, temperature and humidity, and some orifice meter/manometer combinations to monitor the fuel input rates for the main burner, the pilot light and the smoke generator. There are also viewports for visual and camera observations.

Each test begins with the recording of sensor readings to establish baseline values and to verify that the sensors are in working order. Some instruments, notably the gas analyzers in the exhaust hood, produce unusably noisy raw data that are smoothed before recording. Smoothing filters cause the reported values to lag events by 15 to 30 seconds. This lag is considered unimportant under the circumstances. However, these data along with baseline readings are available for post-test data processing. There are also some time lags in the system that should be determined during data processing. One of these is the time needed for hot gases to fill the tops of the rooms down to the level where escape out the doorway starts. Transport times are others, which are the times for effects to move from one location to another.

Smoke is inserted into the otherwise clear fire plume as a visible hot gas tracer. Several smoke sources were tried, the latest being acetylene with incomplete combustion.

The two rooms, an exhaust chamber and a gas collecting hood are inside a building in which stable air conditions are maintained. The exhaust hood contains a thermocouple array and instruments for measuring the percentages of carbon monoxide, carbon dioxide and oxygen in the exhaust gases. These measurements are used as cross-checks on the completeness of burn (through oxygen depletion) of combustibles in the fuel assembly. Experience has shown that these measurements have greater uncertainties than direct metering of the input fuel where a gas fire is used.

The fuel assembly in the present case consists of natural gas for the pilot light and the main burner, and the smoke tracer fuel. Fuel assemblies in later tests will be more complicated, including furniture, pools of combustible liquids, open stacks of wood blocks (called "cribs"), and other possibilities.

There is deliberate redundancy in the instrumentation, meaning that some of the important parameters can be computed in more than one way. This permits selected segments of the test data to be tested for internal consistency. One fruitful check of this kind was for conservation of energy in each of the rooms. As results, the Coefficient of Discharge of the doorways leaving the burn room and the corridor was changed from default values of 1.0 to a value in the 0.68 to 0.75 range, and thermal properties (specific heat, density, conductivity, emissivity) for the burn room floor were corrected.

The heat input rate is cross-checked in several ways. The most direct measure uses the heating value of the fuel and the gas consumption rate. Another method uses oxygen depletion measurements at the exhaust hood. A less direct check involves the conservation of energy calculations in which an error of closure may be partly due to an error in the heat input rate.

Heat absorption through enclosure surfaces is another important factor in the heat balance calculations. The main reliance here is on calculations that use the thermal properties of the materials, accumulated heat absorption to that time in the test, and the local time-varying conditions. Very limited opportunities for cross-checking are provided by the radiometers and by thermocouples on the surface and buried below the surface at some locations.

4.2 Test Procedures

CAM was asked to advise CFR on the "experimental design" with respect to the sequence of tests. The basic recommendations were to randomize among the various fire sizes and to replicate tests in order to provide a statistical basis for estimating variability of the results. Another recommendation was to take and keep data throughout the baseline, pilot-only, main burn and a portion of the cool-down period.

Fire size was the only significant independent parameter. Fire size is described in terms of the heat input rate in kilowatts, kW, as a function of time. Nominal fire sizes are 50kW, 100kW, 300kW and 500kW. The present series of tests employ a step function main burner input. Therefore, these fires are described by the height of this step.

A letter suffix distinguishes different tests at the same input rate--e.g. 100A, 100B, etc. for the 100kW series. The 100kW fire was chosen for the main reference fire, and was re-run more than any of the other choices (other than the pilot light case which was always present).

The active part of a test has four distinct phases separated by step function changes:

1. Establish baseline readings for the sensors--pre-burn conditions. This takes five to ten minutes.

2. Turn on the pilot light and take readings for about five minutes.

3. Turn on the main burner and the smoke generator for ten-fifteen minutes.

4. Turn off the main burner and the smoke. Take readings for another five to ten minutes.

This procedure yields four sets of data:

1. Initial conditions. Ideally, the rooms would be at outside conditions and the air should be still. However, the rooms did not always have time to cool completely from the previous test. Also, there are apparently some minor heat inputs from such sources as floodlights and workers that establish detectable layers and gas flows. The model should be run with these actual initial conditions as well as with nominal valuesto evaluate their effects.

2. Response to a low-level fire, represented by the pilot light phase (about three kW). This phase provides an opportunity to anchor the analyses at the low end of the fire size spectrum.

3. Response to the combined main burner, pilot and smoke generator inputs.

4. Cool-down response, at least for the early phase. These data are not strictly necessary, but they cost very little and might be useful in estimating the dynamic response characteristics of the test facility and the individual rooms.

The aggregate data recording interval is about half an hour, or about 180 ten-second periods. At 150 sensor readings per interval, this leads to about 2700 automated readings per test. Off-line processing generates about 100 channels more of derived results. There are 38 such sets of data from the main test series.

4.3 Data Recording and Processing

CFR has developed a collection of computer routines called SPEEDY that display and interpret the digitized sensor data. SPEEDY and space for the original and derived data files occupy about one megabyte of RAM. Substantial parts of both the SPEEDY and model programs support the presentation of results and are not involved in the calculation of values. Even setting the display segments aside, data interpretation involves an important amount of modeling.

Details of the collection and handling of test data can affect the useful accuracy of the basic data and quantities derived from them, and hence the quality of validation conclusions. Each step in the data reduction process contributes to the accumulated uncertainties and should be examined. Many are relatively inconsequential, but some will be important and irreducable from a practical standpoint. An analysis is recommended to place these error sources in perspective and to identify those places where concentration could be fruitful. Correspondingly, there are places where errors are less important and where useful simplifications may be possible. Of course, the results of such work should be reported so that these lessons will not have to be relearned. Instrument selection and placement, the recorded data and the models and algorithms that use them need to be considered as systems that produce the answers that are the real objectives.

The output sensors produce voltages that are digitized, polled and recorded each ten seconds during a test. Sensor values at the instant of polling are recorded. Variations since the last reading are not available. Thus, the highest frequency of output effects that can be detected is one per twenty seconds.

Baseline values are recorded before the pilot light is lit, but there is no independent determination of initial temperatures or of systematic differences among sensors. Outputs are taken at face value unless gross errors are evident.

The sizes and effects of random errors is another field awaiting systematic attention. Data from each sensor channel are rounded and recorded to an assigned resolution. For example, temperatures are recorded to the nearest whole degree Kelvin, which is considered close enough for the intended use of the data, and less accurate than could be obtained from the thermocouples. Instruments that measure pressure differences through walls have a resolution of five percent of full scale. The equivalent information for other sensor types is not readily available.

4.4 Data Review and Interpretation

The purpose of this step is to find and guide the correction of unanticipated problems. A degree of creativity is useful in devising broad tests with the goal of progressively isolating any perceived causes of difficulty. The formal nature of the step is top-down "analysis", the complement of the building process of "synthesis". The content of this step depends heavily on the particular situation. Rather than attempt to generalize, some examples are provided to illustrate the approach being used here. Problems can arise anywhere from the original concept of the test through data reduction to its eventual interpretation. The tools are a combination of physical reasoning, mathematics, choice of presentation methods and common sense.

The goal is to develop confidence that the events of the test are fairly and accurately represented. Foremost is the detection and reduction of systematic errors. Second is the estimation of random uncertainties. The quantification of these two types of error motivated much of the test planning and instrumentation. Test replication addresses random uncertainties.

Measuring more than the absolute minimum number of parameters permits the testing of submodels and supports expert study of results for their combined reasonableness. For example, there are some thermocouples buried at various depths in the wall materials as aids in checking enclosure heat absorption models.

One pair of generic tests is the application fo Conservation of Mass and Conservation of Energy laws. In attempting to reconcile terms in a burn room heat balance, it became clear that the thermal characteristics of the burn room floor material were incorrect. Burn room heat balances for at least one test now reconcile within ten percent. A mass balance test led to a change in doorway coefficients of discharge.

Plots of raw and derived parameters have proven to be valuable early tools for experts to detect and diagnose unsuspected problems. For example, are the slopes, asymptotes and relationships to other data reasonable? Guidance on noise levels can also be obtained. Tests can be made on raw and processed data. Unexpected aberrations are clues to a range of problems, some of which have been detected in this manner.

Plots do not replace more formal statistical tests, but can obviate them in some cases. On the negative side, formal tests tend to produce summary results and in the process discard some of the information content in the basic data that plots would retain.

Plots of some temperatures against time led to amendment of one of the processing algorithms. The initial algorithm choice used zones bounded at the mid-points between sensors and assumed that the the temperature and flow at each sensor applied uniformly over its zone. Some erratic and unbelievable results occurred where the layer boundary oscillated above and below one of the sensors. Figure 3 is a time plot of the average temperature of the gases leaving the burn room. The jagged pattern part way up on the rise and as the temperature leveled off is strongly counter-intuitive. This led to a review of the algorithm used in SPEEDY to derive this time series.

The assumption of linear changes between sensors suppressed some of the erratic behavior seen in Figure 3, with unevaluated effects on the other objectives. The shortage of sensors in the upper layer limits possibilities for higher-order curvilinear fits.

Estimation of the mass flow profile through a doorway provides examples where there may be opportunities to improve the combined instrumentation and computation plan. The issues are the selection and placement of the sensors and the design of the data reduction algorithm.

The mass flow rate at a given elevation in a doorway is proportional to the square root of (T x P) where T is absolute temperature and P is pressure difference through a wall at the door. Therefore, a percentage error in pressure is just as important to the answer as the same percentage error in temperature. Since relative errors are inherently smaller for temperature

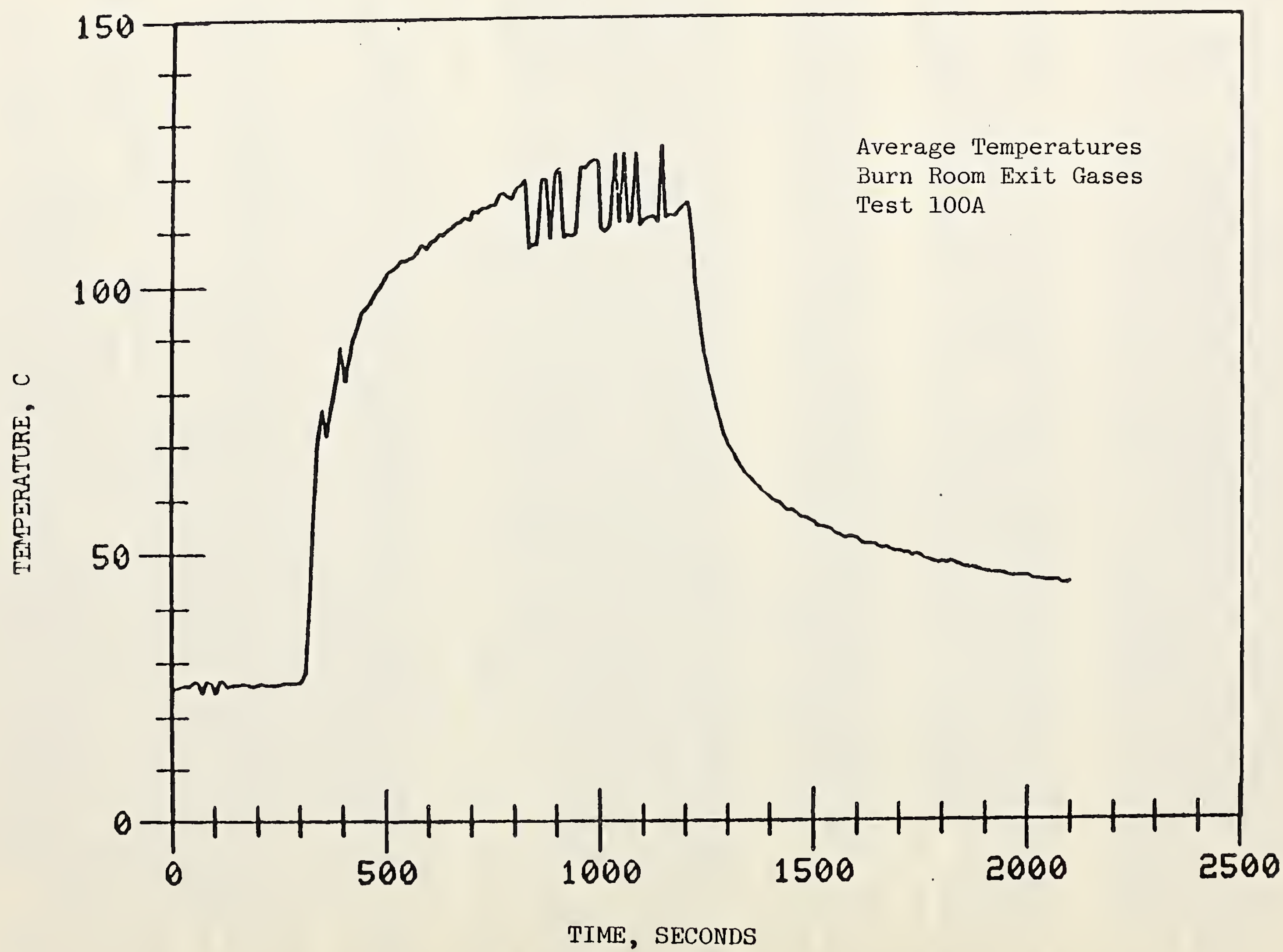

FIGURE 3.
TEMPERATURE VS TIME, TEST 100A

than for pressure, designs placing greater reliance on temperature are preferred. Additionally, the flow and hence the pressure difference at the top of the door is away from the fire, say out, and in at the bottom of the door. Somewhere between the top and the bottom, the pressure difference must pass through zero whereas the absolute temperature is large and never zero throughout such a traverse. The possibility of directional errors near this crossover level also enter the problem.

There are three competing sensor tree arrangements, with a flow computation at each level of the tree. One method uses vertical trees of paired thermocouples and pressure sensors. The second method uses a single temperature taken at some advantageous height, together with the vertical array of pressures. The third method reverses the roles of the pressure and temperature sensors, using a single pressure and several temperatures.

The third method is preferred, with the pressure is measured at a level where it is large, and hence has a small relative error. Also, if the original complement of pressure probes remains available, they could all be located at the chosen elevation and their results averaged to reduce the resultant uncertainty relative to that from a single probe.

There are also curve-fitting choices to be made in estimating the flow profile, given the flow estimates at small number (ten or less) of specific elevations. The compromise is among estimating the height of the steep thermal gradient at the boundary between the hot and cold layers, following the profiles in these distinct zones, and satisfying conservation of mass and energy requirements.

5. VALIDATION ACTIVITIES

Most of the work so far has been on technical validation -gaining confidence in the numerical quality of the results. Progress on user validation has been slow. Documentation is sparse and not yet integrated. Several large steps remain before trials by outside users could start. As of October 1984, a major round of improvements to SPEEDY has been accomplished, but tests of these changes remain to be made. Comparison studies and the development of uncertainty statistics have not begun.

Processes for the graphing of test and model results have been demonstrated with synthetic inputs using DATAPLOT [4] on the UNIVAC. The installation of selected portions of DATAPLOT on the Perkin-Elmer 3242 at CFR is being explored.

Figure 3 (introduced earlier) is one such product. Figure 4 is another temperature plot involving Test 100A data. The purpose here is to illustrate how model and test data might look together.

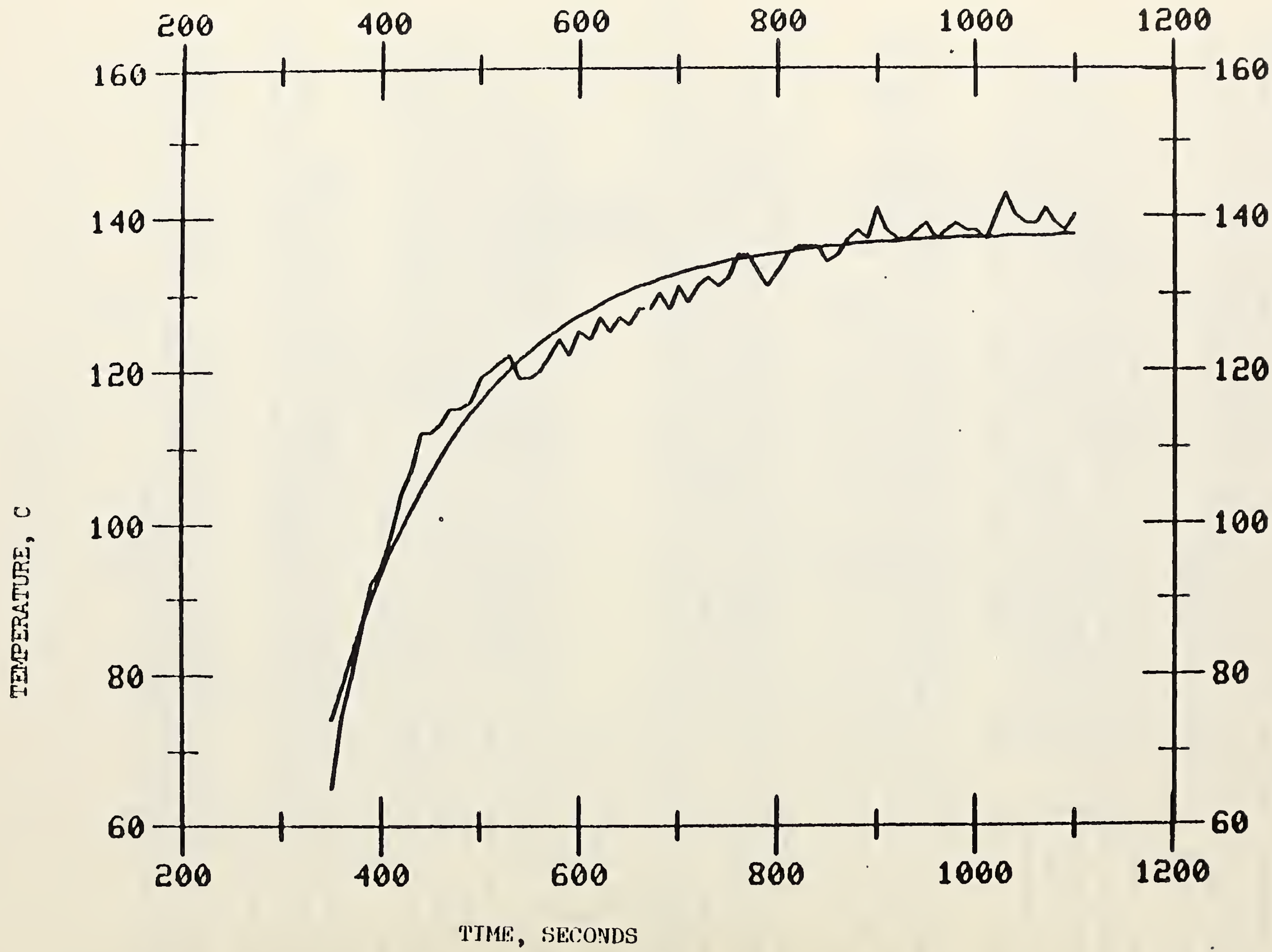

Figure 4. Temperature vs Time

Test and Regression Results

The effect of polling the sensors each ten seconds is seen in the short straight-line segments in one of the curves. Model results were not available at the time, so a stand-in curve was generated by DATAPLOT. It is a four-constant regression fit of Temperature = A + B*EXP(C*(Time -D)) to a selected segment of temperature data. The regression curve is inherently smooth, whereas the model results might not be. Nevertheless, much perspective can be gained from such a plot.

° The main trends of the curves cross more than once, as should be expected from the multiple-constant regression formula that was used.

° The sensor temperatures are erratic, with limited departure from the main trend.

° A timing error between the two curves would translate into an apparent temperature difference. This effect is greatest where the slopes of these curves is greatest.

° A small handful of constants "explained" most of the behavior of the test results. This leads to the suspicion that many of the test parameters might be combined into more concise descriptors before applying the model and that several of the inputs have a minor effect on the outcome. This does not mean that they can be ignored, but their effects may be masked by uncertainties in some of the more important inputs. This motivates the quantification of uncertainties in the inputs and the propagation of these effects through the computations.

Figure 5 is another DATAPLOT product in which the residual differences between the two curves in Figure 4 are plotted as connected points. Figure 6 is a similar plot where the discrete nature (i.e. once per ten seconds) of the test and model data is recognized without implying any assumptions about the behavior between polling times. The same kinds of information are in Figures 5 and 6, although Figure 6 came from a different data set. Both versions may be useful since different observers may see different things in the two plots.

° The vertical scale can be expanded since we are now dealing with differences between similar curves. Crossings of the original curves now appear as crossings of the zero difference axis. If the sizes of these residuals are small enough as they stand, the user may decide not to look any deeper.

° There is no fixed bias term in the errors (due to the way the regression curves were generated), but there is some other systematic character (called "structure") that might be reduced with a more refined model, changes in instrumentation or changes in SPEEDY. In Figure 5, the relatively high amplitude of the residuals up to about 550 seconds has been discussed above. Thereafter the amplitudes are smaller, and a generally positive trend can be detected. The actions taken to deal with these two observations may differ.

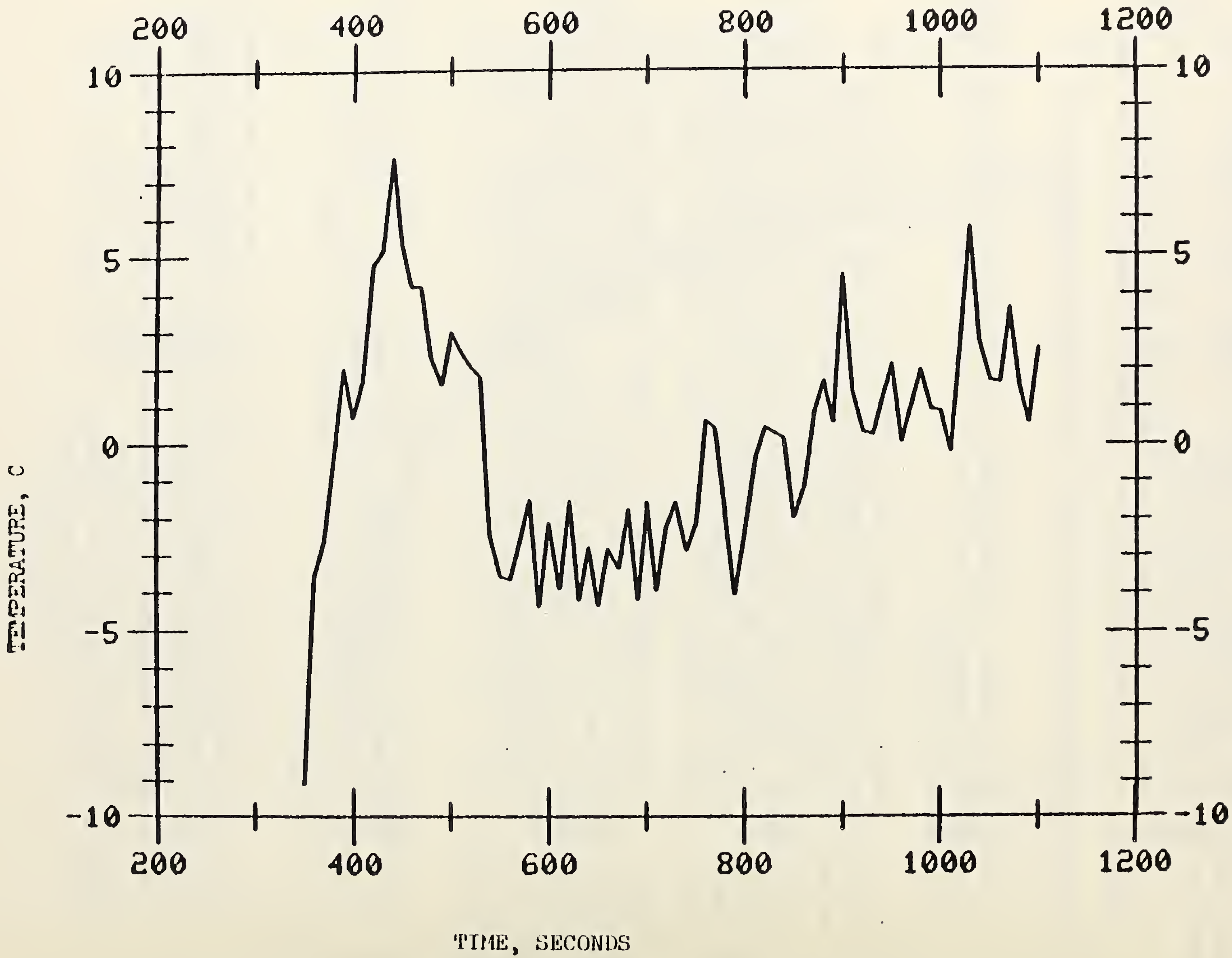

Figure 5. Residual Difference Curve

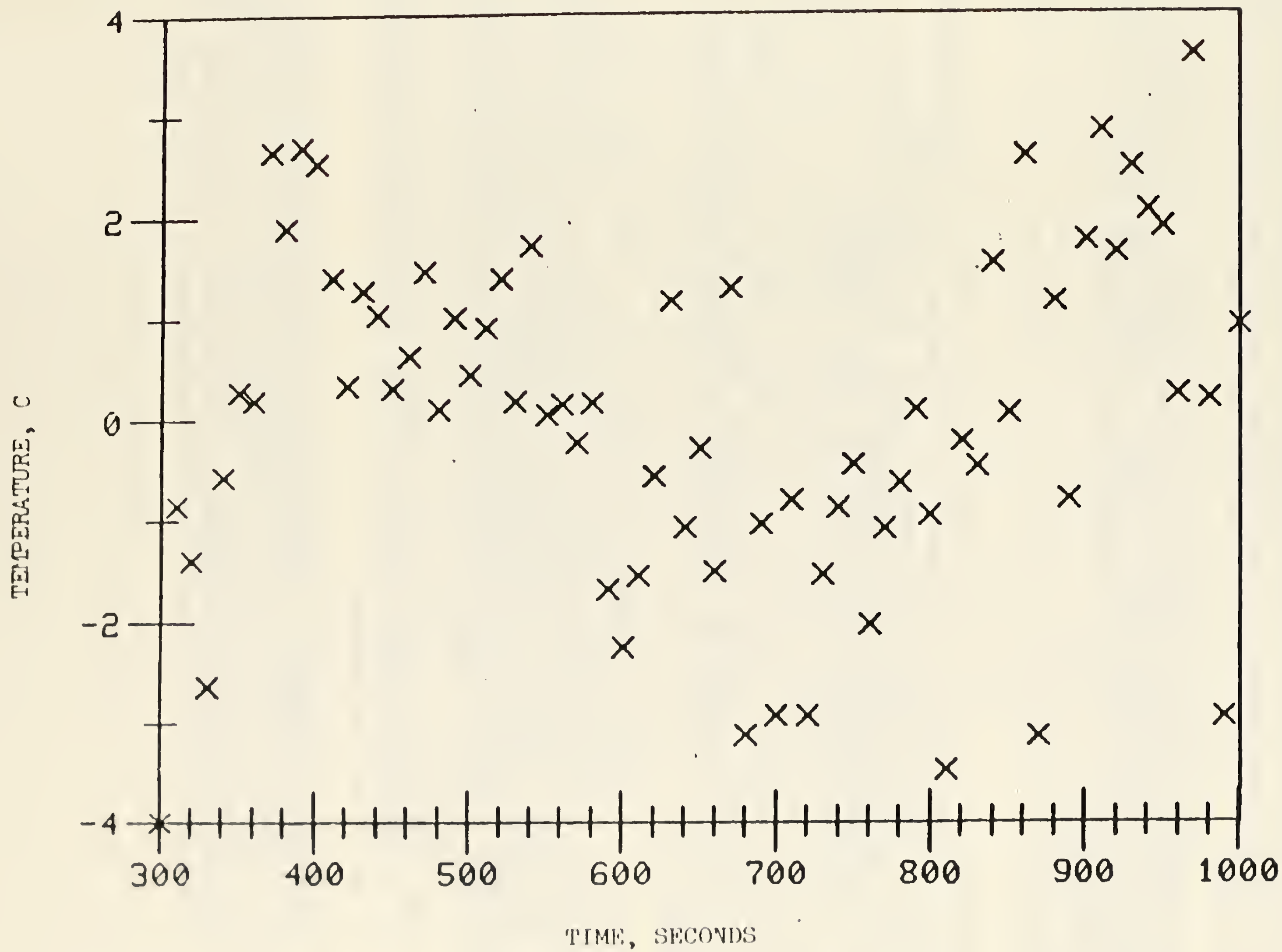

FIGURE 6. RESIDUAL DIFFERENCE POINTS

° The errors seem to have several frequency components, each with its own amplitude. The low frequency is associated with the zero crossings (which is an artifact of the regression) and the higher ones are not yet explained. The user might isolate the components of most interest by selectively smoothing and filtering these residuals.

° Residual difference plots can be devised for other types of tests such as for conservation of energy and mass. Essentially, the residuals are computed and plotted for any combination of terms that should be zero. Bias effects are present if the residuals are not balanced around the zero axis.

° Some parameters such as temperatures from sensors on a thermocouple tree should be systematically related. Plotting them together on the same time graph as well as cross-plotting one against another can be instructive.

6. OBSERVATIONS

° Validation of the test data turned out to be a substantial and important task in preparation for the numerical validation of the model. The reason is that residual differences between the model and the test outputs are needed to tell a potential user how "good" the model is. Such statements depend both on test and model results.

° There is as much or more software and modeling needed to translate the raw sensor readings from the tests into terms comparable with model outputs as there is in the model itself. Errors in interpreting the raw data can be as serious and difficult to detect as problems in the model.

° The main deficiency for validation purposes was in the documentation area, both from the experimental and modeling sides. Good documentation is hard to produce and takes more time and work than is usually allowed in the plans. Several versions may be appropriate as the project progresses and the audience moves from project personnel to the user of a released package. Types of documentation that are needed include:

User information -description in non-ADP terms of functions performed by the software so that a potential user can determine where and how to use it and get answers.

Operations information - description of the software and machine environment needed to load the program and data and run the program.

Modeling concepts -the mathematical and physical concepts that are to be supported, with formulas that are recognizable in the program.

Input data requirements -sources, definitions, dimensions, preparation, symbols.

Program documentation - listing, flow diagrams, other information to enable understanding in support of program maintenance.

Experimental Reference -information needed to replicate the test facility and the tests.

° A free-form project log book has turned out to be quite useful, rather than depending on the memories of the participants regarding the experimental setup, procedures and any notable problems or events in a given case. Such information can help in diagnosing problems in the analysis, or even much later when the project is over and key personnel are scattered. Records are also important in preparing documentation, if only to improve team coordination. Attention needs to be given to the uncertainties in measurements, particularly as they appear on the data tape or other basic record from which analyses begin.

7. RECOMMENDATIONS

° A work breakdown structure of tasks and an associated schedule was developed and provided by CAM. Some equivalent written plan is recommended, even though it may be revised as the project proceeds. People need to know who is responsible for what and when the results are expected.

° Arrange for the development and printout of intermediate and final variables in the model that correspond to test results. These should be selected in consultation with the testers with the idea of checking out portions of the model in the inevitable debugging process.

° Institute the project position of "technical coordinator" who will have charge of interface descriptions, supervision of documentation and generally to make sure that the various team groups understand what each other is doing and that the work is technically compatible. This person would also have schedule and cost responsibilities. Identifiable functions on the present project include modeler, test designer and manager, facility builder, tester, test data programmer, data collector.

° Develop a computation and display of moving estimates of standard deviations or their equivalent directly from the data streams. The purpose here is to develop variability statistics as a function of time that are not too heavily weighted at other times during the test.

REFERENCES

1. Baum, H. R. et al, "Direct Computation of Fire Induced Flow and Smoke Coagulation"; NBS, Presented at the Combustion Institute, Eastern States Section "Chemical and Physical Processes in Combustion.", 1981.

2. Baum, H. R. et al, "Prediction of Heat and Smoke Movement in Enclosure Fires"; NBS, Presented at the Eighth International Symposium on Problems of Automatic Fire Detection, 1982.

3. Cooper, L. Y., Stroup, D. W., "Calculating Available Safe Egress Time (ASET) - A Computer Program and User's Guide"; NBSIR 82-2578, National Bureau of Standards, 1982.

4. Filliben, J. J., "DATAPLOT (tm) (Version 84/7)" (available through NTIS), Nat. Bur. Stds., 1984.

5. Francis, R. L., Saunders, P. B., "EVACNET: Prototype Network Optimization Models for Building Evacuation"; NBSIR 79-1738, Nat. Bur. Stds., 1979.

6. Gass, S.I., "Computer Model Documentation: A Review and an Approach"; NBS Special Publication 500-39, Nat. Bur. Stds., 1979.

7. Gass, S.I. (ed), "Validation and Assessment Issues of Energy Models"; NBS Special Publication 569, Nat. Bur. Stds., 1979.

8. Gass, S.I. et al, "Interim Report on Model Assessment Methodology: Documentation Assessment"; NBSIR 80-1971, Nat. Bur. Stds., 1980.

9. Gass, S.I. et al, "The NBS Energy Model Assessment Project: Summary and Overview"; NBSIR 80-2128, Nat. Bur. Stds., 1980.

10. Gass, S.I., "Oil and Gas Supply Modeling"; NBS Special Publication 631, Nat. Bur. Stds., 1980.

11. Jones, W. W., "A Review of Compartment Fire Models"; NBSIR 83-2684, Nat. Bur. Stds., 1983.

12. Jones, W.W., "A Model for the Transport of Fire, Smoke and Toxic Gases (FAST)"; NBSIR 84-2934, Nat. Bur. Stds., 1984.

13. Quintiere, J. G., DenBraven, K., "Some Theoretical Aspect of Fire Induced Flows Through Doorways in a Room-Corridor Scale Model"; NBSIR 78-1512, Nat. Bur. Stds., 1978.

14. Rockett, J. A., "Modeling the NBS Mattress Tests with the Harvard Mark V Fire Simulation"; NBSIR 81-2440, Nat. Bur. Stds., 1982.

15. Saunders, P.B. (ed), "Selected Assessment Strategies Applied to Short-Term Energy Models"; NBSIR 83-2672, National Bureau of Standards, 1983.

16. Stahl, F. I. et al, "Time-Based Capabilities of Occupant to Escape Fires in Public Buildings: A Review of Code Provisions and Technical Literature"; NBSIR 82-2480, Nat. Bur. Stds., 1982.

17. Tanaka, T., "A Model of Multiroom Fire Spread"; NBSIR 83-2718, Nat. Bur. Stds., 1983.

NBS-114A (REV. 2-80)

U.S. DEPT. OF COMM. BIBLIOGRAPHIC DATA SHEET (See instructions)	1. PUBLICATION OR REPORT NO. NBSIR 85-3154-1	2. Performing Organ. Report No.	3. Publication Date July 1985

4. TITLE AND SUBTITLE

Applied Model Validation

5. AUTHOR(S)

Alan D. Davies

6. PERFORMING ORGANIZATION (If joint or other than NBS, see instructions)	7. Contract/Grant No.
NATIONAL BUREAU OF STANDARDS DEPARTMENT OF COMMERCE WASHINGTON, D.C. 20234	8. Type of Report & Period Covered

9. SPONSORING ORGANIZATION NAME AND COMPLETE ADDRESS (Street, City, State, ZIP)

Center for Fire Research
National Bureau of Standards
Gaithersburg, MD 20899

10. SUPPLEMENTARY NOTES

☐ Document describes a computer program; SF-185, FIPS Software Summary, is attached.

11. ABSTRACT (A 200-word or less factual summary of most significant information. If document includes a significant bibliography or literature survey, mention it here)

This is a progress report of an applied model validation case study. The subject model is "Transport of Fire, Smoke and Toxic Gases (FAST)" by W. W. Jones of the National Bureau of Standards, Center for Fire Research. Products from a fire in a "burn room" exit through a connected corridor to outdoors. Cooler counterflow air in a lower layer feeds the fire. The model predicts corridor layer temperatures and thicknesses vs. time, given enclosure, fire and ambient specifications. Data have been collected from 38 tests using several fire sizes, but have not been reduced. Corresponding model results, and model and test documentation are yet to come.

Considerable modeling and calculation is needed to convert instrument readings to test results comparable with model outputs so that residual differences may be determined. Test results as well as model results must be validated, and test result uncertainties estimated so that they are not unfairly attributed to the model.

12. KEY WORDS (Six to twelve entries; alphabetical order; capitalize only proper names; and separate key words by semicolons)

Data editing; experiments; fire models; fire research; residual difference; validation

13. AVAILABILITY	14. NO. OF PRINTED PAGES
[X] Unlimited ☐ For Official Distribution. Do Not Release to NTIS ☐ Order From Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402. [X] Order From National Technical Information Service (NTIS), Springfield, VA. 22161	31 15. Price $8.50

USCOMM-DC 6043-P80